DOCUMENT RESUME

ED 118 588 TM 005 071

TITLE Measuring Student Achievement in Home Economics.
INSTITUTION New York State Education Dept., Albany. Bureau of Elementary and Secondary Educational Testing.; New York State Education Dept., Albany. Bureau of Home Economics Education.
PUB DATE 75
NOTE 56p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage
DESCRIPTORS *Academic Achievement; Elementary Secondary Education; Guidelines; Home Economics; *Home Economics Education; *Measurement; Student Evaluation; Student Testing; *Test Construction

ABSTRACT

Many home economics teachers have experienced difficulty in developing good classroom tests and have expressed a need for help. This publication was prepared to provide guidance in the construction and use of a variety of measurement devices. Section 1 discusses the role of measurement in the instructional process including why measure, what to measure, and how to measure. Section 2 deals with preparing instruments used to assign grades. Advice is given on preparing a test blueprint, choosing an item type (essay, completion, true-false, multiple-choice, and matching), and measuring cognitive and psychomotor learning outcomes. Section 3 discusses preparing instruments used to evaluate and plan instruction related to cognitive, psychomotor, and affective learning outcomes. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED 118 558

MEASURING STUDENT ACHIEVEMENT IN HOME ECONOMICS

DEC 5 1975

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM005 071

The University of the State of New York
THE STATE EDUCATION DEPARTMENT
Bureau of Elementary and Secondary Educational Testing
Albany, New York 12234

MEASURING STUDENT ACHIEVEMENT

in

HOME ECONOMICS

The University of the State of New York
THE STATE EDUCATION DEPARTMENT
Bureau of Home Economics Education
Bureau of Elementary and Secondary Educational Testing
Albany, 1975

THE UNIVERSITY OF THE STATE OF NEW YORK

Regents of The University *(with years when terms expire)*

1981	THEODORE M. BLACK, A.B., Litt.D., LL.D., Pd.D., D.C.L. *Chancellor*	Sands Point
1987	CARL H. PFORZHEIMER, JR., A.B., M.B.A., D.C.S., H.H.D. *Vice Chancellor*	Purchase
1978	ALEXANDER J. ALLAN, JR., LL.D., Litt.D.	Troy
1980	JOSEPH T. KING, LL.B.	Shelter Island
1981	JOSEPH C. INDELICATO, M.D.	Brooklyn
1979	FRANCIS W. MCGINLEY, B.S., J.D., LL.D.	Glens Falls
1986	KENNETH B. CLARK, A.B., M.S., Ph.D., LL.D., L.H.D., D.Sc.	Hastings on Hudson
1983	HAROLD E. NEWCOMB, B.A.	Owego
1988	WILLARD A. GENRICH, LL.B., L.H.D., LL.D.	Buffalo
1982	EMLYN I. GRIFFITH, A.B., J.D.	Rome
1977	GENEVIEVE S. KLEIN, B.S., M.A.	Bayside
1981	WILLIAM JOVANOVICH, A.B., LL.D., Litt.D., L.H.D.	Briarcliff Manor
1976	MARY ALICE KENDALL, B.S.	Irondequoit
1984	JORGE L. BATISTA, B.A., J.D.	Bronx
1982	LOUIS E. YAVNER, LL.B.	New York

President of The University and Commissioner of Education
EWALD B. NYQUIST

Executive Deputy Commissioner of Education
GORDON M. AMBACH

FOREWORD

The measurement of student achievement is an integral part of the teaching-learning process. In order to effectively guide student growth, a teacher must know what objectives have been mastered and what objectives have yet to be attained. This requires the preparation of valid and reliable measuring instruments which may be used to determine readiness for new learnings, pinpoint areas for remedial work, and provide guidance for improving instruction.

Many home economics teachers have experienced difficulty in developing good classroom tests and have expressed a need for help. Therefore, this publication was prepared to provide guidance in the construction and use of a variety of measurement devices. The development of this publication was completed under the direction of Elizabeth A. Brown, Chief of the Bureau of Home Economics Education. Assistance relating to content was provided by Carol Jabonaski, Associate in the Bureau of Home Economics Education. The manuscript was prepared and coordinated by Kathleen Scott, Consultant, and Kenneth Ormiston, Associate in the Bureau of Elementary and Secondary Educational Testing.

Winsor A. Lott, Chief
Bureau of Elementary and Secondary Educational Testing

Victor A. Taber, Director
Division of Educational Testing

Contents

	PAGE
I. The Role of Measurement in the Instructional Process	1
A. Why Measure?	1
B. What To Measure	2
C. How To Measure	3
II. Preparing Instruments Used To Assign Student Grades	5
A. Preparing a Test Blueprint	5
B. Choosing an Item Type	9
1. Essay Questions	10
2. Completion Items	10
3. True-False Items	11
4. Multiple-Choice Items	11
5. Matching Questions	12
C. Measuring Achievement To Assign Grades: Cognitive Learning Outcomes	13
1. Essay Questions	13
2. Completion Items	16
3. True-False Items	18
4. Multiple-Choice Items	20
5. Matching Items	24
D. Measuring Achievement To Assign Grades: Psychomotor Learning Outcomes	27
1. Introduction	27
2. Performance Test Construction Procedures	28
3. Sample Performance Test Exercise	29
III. Preparing Instruments Used To Evaluate and Plan Instruction	34
A. Evaluating and Planning Instruction: Cognitive Learning Outcomes	34
B. Evaluating and Planning Instruction: Psychomotor Learning Outcomes	34
C. Evaluating and Planning Instruction: Affective Learning Outcomes	38
1. Introduction	38
2. Questionnaires and Inventories	38
3. Logs and Diaries	43
4. Anecdotal Records	47
References	50

I. THE ROLE OF MEASUREMENT IN THE INSTRUCTIONAL PROCESS

A. Why Measure?

Imagine, if you will, a world without any kind of measurement: no clocks, no yardsticks, no measuring cups, no thermometers, no scales . . . no measuring instruments of any kind. What confusion! It might be fun for a day or two, but most of us would soon find such a world just too chaotic to enjoy.

As teachers, we would probably feel the same way about an instructional program without measurement. It would seem nice at first, but before long we would be hopelessly confused and frustrated. Both student and teacher need to know when they have reached their goals and when they have fallen short. In addition, measurement plays a vital role in the effective management of schools and school systems as a basis for evaluating and planning school programs. Moreover, there is a growing concern for accountability in education. Those who support our educational enterprises want to know how well the job is being done. The questions which are being raised cannot be answered without the information obtained through measurement.

This publication is intended for teachers and deals with measurement in the classroom. A few reasons why measurement is important are listed below.

1. To Plan Instruction. If achievement is measured before instruction, instruction can be tailored to meet the needs of the students. In addition, the students will better understand the specific objectives for instruction.

2. To Motivate Students. Most students will exert a greater effort to learn when they know that their achievement will be measured from time to time.

3. To Evaluate Instruction. The extent to which students obtain an objective is one indication of the effectiveness of instruction.

4. To Assign Grades. Perhaps the most obvious reason for measuring achievement is to assign grades which are a fair and accurate measure of student growth.

B. What To Measure

The main goal of the measurement process is to determine the extent to which students have attained the desired objectives. Each segment of the courses of study contained in the State syllabus and each module in the updated curriculum package for home economics education includes a list of "behavioral outcomes." These outcomes describe the behaviors which students should exhibit at the conclusion of instruction.

These lists of outcomes describe a variety of behaviors, but all of them can be included under one of three headings or domains: cognitive, psychomotor, and affective. A brief explanation of each of these three categories of behaviors follows.

Behaviors within the cognitive domain "... deal with the recall or recognition of knowledge and the development of intellectual abilities and skills."[1] The knowledge outcomes relate to the psychological processes of remembering, while the intellectual abilities and skills outcomes relate to the mental processes of organizing and reorganizing material to achieve a particular purpose.

Behaviors within the psychomotor domain include manipulative skills. Behaviors within the affective domain involve "... changes in interest, attitudes, and values, and the development of appreciations and adequate adjustment."[2]

Following are a few examples of the kinds of outcomes that would be included in each of the three domains.

1. Cognitive Learning Outcomes:

- Knows the contribution of nutrients to good health.
- Recognizes situations which might produce safety hazards for a child.
- Is able to identify factors important to consider when choosing a location for living.
- Understands how family, culture, and community can influence food habits.
- Is able to analyze relationships between family health, safety, harmony, and the maintenance and upkeep of living space.
- Is able to integrate the many costs related to housing into a workable financial plan for families of varying circumstances.

[1]Bloom, B. S., ed., *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1: Cognitive Domain* (New York: McKay, 1956), p. 7.

[2]Ibid.

2. **Psychomotor Learning Outcomes:**
   - Is able to use appropriate first aid procedures when emergencies arise.
   - Is able to make a well-constructed garment.
   - Is able to plan, prepare, and serve appropriate foods for various types of entertaining.
   - Is able to care for the clothing of infants and young children.

3. **Affective Learning Outcomes:**
   - Shows interest in preparing self for a career.
   - Accepts each child as an individual.
   - Enjoys assisting in planning, preparing, and serving family meals.
   - Takes pride in own cultural heritage.
   - Feels some security in own adolescent transition.

All three types of outcomes are important, and the teacher should attempt to measure the extent to which all students have achieved in each domain. The teacher must be aware, however, that the techniques used to measure achievement for each of the three types of outcomes differ greatly.

C. How To Measure

Different types of instruments can be used to obtain information about student achievement. Paper and pencil tests utilizing objective (multiple-choice, completion, true-false, matching) and essay questions, performance tests, questionnaires, inventories, logs, diaries, and anecdotal records can all provide valuable information.

When measuring pupil progress in order to assign grades, only the cognitive and psychomotor outcomes should be measured. It is difficult to obtain a reliable measure of student achievement of the affective outcomes. What is most often measured is the students' knowledge of the attitudes considered acceptable, not the students' actual attitudes. This is especially true if the students know that the test results will be used to assign grades. Also, it is often impossible to construct test questions that have a single correct answer when measuring the affective outcomes. Each question may have several correct answers depending on the students' values. Students should not be penalized if they express interests, attitudes, or values which are different from those of the teacher. Paper and pencil tests utilizing objective and essay questions, and performance tests are used to measure achievement of the

cognitive and psychomotor outcomes. The second section of this booklet will contain suggestions for preparing these types of instruments.

Additional reasons for measuring achievement are to plan instruction and to judge the effectiveness of instruction as well as to help students to see their progress. The third section of the booklet will contain suggestions for preparing instruments that can be used for these purposes.

Following is a table that shows which types of instruments may be used by the teacher to assign grades and which types may be used to plan and evaluate instruction for each of the three types of outcomes.

Table 1

Selection and Use of Instruments

Behavioral Learning Outcomes	Instruments That Can Be Used To Assign Grades	Instruments That Can Be Used To Evaluate and Plan Instruction
Cognitive (knowledge and intellectual abilities and skills)	Paper and pencil tests utilizing objective and essay questions	Paper and pencil tests utilizing objective and essay questions
Psychomotor (manipulative skills)	Performance tests (scored by teacher only)	Performance tests (scored by teacher, other students, or students themselves)
Affective (attitudes, values, and interests)	None	Questionnaires, inventories, logs, diaries, and anecdotal records

II. PREPARING INSTRUMENTS USED TO ASSIGN STUDENT GRADES

A. Preparing a Test Blueprint

When preparing any test instrument, the first step is to prepare a test blueprint. In preparing the blueprint, the teacher must decide what cognitive and psychomotor outcomes are to be measured by the test and what their relative weights should be. These factors must be considered if the teacher is to construct a test which will provide a valid picture of pupil progress.

There is no single type of blueprint that is best for all tests. For a brief quiz, a listing of outcomes and their relative weights is sufficient. A semester or final examination, however, requires a much more comprehensive blueprint.

One sample blueprint for a semester test in home economics is presented in table 2. The major steps followed in constructing this blueprint are listed below:

1) Listing of the modules and behavioral outcomes to be included in the test and the amount of instructional time (in hours) that was required for each module.
2) Determination of the percentage of the total test that should be allocated to each module (last column). Each percentage is determined by dividing the instructional time required for each module by the total instructional time.
3) Division of the percentage allocated to each module (as obtained in step 2) into cognitive and psychomotor outcome percentages.

It should be noted that the weighting (step 3) between the cognitive and psychomotor columns would probably vary from teacher to teacher and even from class to class. The main considerations that would affect this weighting are the stated outcomes of each module and the teacher's judgment as to which outcomes are most important.

When the test blueprint has been completed, the teacher must then decide which outcomes are to be measured for each module included in the blueprint. In a semester test, the teacher cannot always measure student progress toward all outcomes for all the modules included in the blueprint. Therefore, the teacher must select those cognitive and/or psychomotor outcomes which are believed to be most important.

Table 2

TEST BLUEPRINT FOR A SEMESTER TEST IN HOME ECONOMICS

Subject Area	Module Titles and Behavioral Outcomes	Number of Hours*	Behavioral Outcomes: Cognitive	Psychomotor	Totals
1 Human Development	a) Make Time Work for You • Develops a helpful plan for more efficient use of time in relation to individual goals. • Utilizes developed plan to achieve immediate and long-term goals. • Examines various careers for ways time functions in meeting job responsibilities.	3	5%	—	5%
	b) Jobs That Deal With People • Identifies home economics-related jobs that deal with people. • Examines a job of special interest in home economics-related occupation dealing with people, delineating responsibilities, education, and skills required.	5	7	—	7
	c) Pride in Family Traditions • Describes traditions and customs of different cultures. • Shares with classmates experiences related to own family traditions and customs or those of others they know.	5	7	—	7
	d) Dear HOPE • Identifies personal problems typical of teens. • Applies problem-solving techniques to personal problems of self and others.	5	7	—	7

2. Housing, Furniture, and Equipment	a) Running Household Cleaning Equipment • Uses owner's manuals to learn how to operate household cleaning equipment. • Uses household cleaning equipment properly.	4	3	3	6
	b) Selecting Accessories for Your Room • Describes factors to consider when selecting accessories for own room. • Selects accessories for own room.	5	7	—	7
	c) Spic and Span • Reduces noise pollution in the kitchen. • Uses procedures that minimize air pollution in the kitchen. • Practices recommended procedures for cutting down on trash.	5	4	3	7
3. Food and Nutrition	a) Plain and Fancy Table Setting • Uses a menu to determine appropriate ways of setting the table and serving various meals. • Creates attractive table settings. • Utilizes attractive and inexpensive centerpieces. • Develops skill and confidence in serving foods.	3	2	3	5
	b) Being an Informed Grocery Shopper • Prepares an adequate food shopping list. • Compares foods in various forms and selects the best for the situation. • Identifies types of food stores and services provided.	5	7	—	7
	c) How Foods Work for You • Identifies the nutrients, their uses in the body, the food sources, and diseases caused by nutritional deficiencies. • Plans menus using essential nutrients.	5	7	—	7

Subject Area	Module Titles and Behavioral Outcomes	Number of Hours*	Behavioral Outcomes: Cognitive	Psychomotor	Totals
	d) Creative Cookies • Identifies the six basic types of cookies. • Creates varied and interesting cookies. • Plans for adequate storage of cookies. • Uses appropriate ways of serving cookies.	5	5	2	7
4. Clothing and Textiles	a) Measuring for Correct Pattern Size • Uses correct techniques for measuring body for patterns. • Analyzes body measurements in relation to pattern sizes. • Uses correct techniques to make simple pattern alterations.	2	1	2	3
	b) Size Up Your Style • Identifies clothing styles suited to various personalities and figure types. • Selects clothing styles suited to own personality and figure type.	3	5	--	5
	c) Sew and Go • Selects a pattern appropriate to experience and skill. • Uses appropriate techniques in constructing a garment. • Identifies acceptable standards in garment construction.	15	12	8	20
	Totals:	70 hrs.	79%	21%	100%

*Amount of instructional time required for this module.

B. Choosing an Item Type

After the blueprint has been prepared, the next step is to select the type of test items that will most effectively measure student achievement on each behavioral objective.[3] The teacher should always be sure that the test item chosen is actually measuring the same type of behavior as is described by the objective. Many times a teacher will construct a test item for an objective that does not measure the behavior described by the objective. An example is shown below.

Example: (Objective)	The student uses all of the safety procedures demonstrated in class when peeling and slicing foods.
(Inappropriate Test Item)	List all of the safety procedures demonstrated in class that should be followed when peeling and slicing foods.

The objective shown above states that the student should be able to use all of the safety procedures demonstrated in class when peeling and slicing foods, and is a psychomotor objective. However, the test item asks the student to only list the appropriate safety procedures, not actually use them while peeling and slicing foods. A more appropriate test item for the objective given above is as follows:

(Appropriate Test Item)	Using the materials provided and all of the safety procedures demonstrated in class, peel and slice a carrot crosswise.

The types of test items most commonly used when measuring student achievement of the cognitive objectives are essay questions, completion items, true-false items, multiple-choice items and matching items; performance tests are used to measure student achievement of psychomotor objectives.

A brief discussion of the advantages and disadvantages of each item type used to measure achievement of the cognitive objectives follows. The construction of each item type is covered in part C and the construction of performance tests is covered in part D of this section.

---

[3] *Homemaking-Family Living, Curriculum Planning Guidelines*, Levels I and II, Grades 5-8. New York State Education Department, 1974, p. 5.

1. Essay Questions

The fundamental contribution of the essay question is that it requires students to develop an answer from their own background or experience, without the benefit of suggested possibilities or alternatives, and to express that answer in their own words. When the teacher wishes to test these skills, the use of essay rather than objective questions is indicated. By requiring students to present evidence, to evaluate, to analyze, and to solve new problems or approach problems in a new way, essay questions can serve to measure some higher level abilities and thus contribute to knowledge about the students.

The primary difficulty in the use of essay questions is achieving reliability of test scores. One source of unreliability in the essay question is subjectivity of rating. Considerable variation of assigned marks has been found, not only between one teacher and another, but in ratings by the same teacher from day to day. Usually, extreme cases of subjectivity of rating can be ascribed to failure to use a proper key or to improper rating procedures, but the point remains that the rating of essay answers is inherently a matter of subjective judgment.

A second source of unreliability is possible inadequacy of sampling. When a test includes relatively few essay questions, the sampling of content in the test is likely to be quite narrow. A student who is strong in one area and weak in another will receive a high score if the test happens to include an essay in the first area and a low score if it happens to include, instead, an essay in the second area.

This does not mean that the teacher should avoid the use of essay questions. It does mean, however, that when essay questions are used, special care should be taken to achieve reasonable reliability of test scores. In addition, it would seem advisable for the teacher to reserve the use of essay questions for measuring outcomes that objective items are not believed to measure satisfactorily.

2. Completion Items

The completion item combines the recall skills needed to answer the essay question with the objectivity of rating achieved by a true-false or multiple-choice item. If completion items are well constructed, the answers are definite and specific and the rating can be done quickly and accurately. Wide sampling of content is possible, since a relatively large number of items can be answered in a short period of time.

The main disadvantage of completion items is that it is very difficult to construct items measuring the higher levels of intellectual skills so that the items have only one correct answer. Teachers are often surprised by

the variety of correct answers given by students for an item which the teacher thought had only one correct answer. As the number of possible responses to an item increases, scoring becomes more difficult and more time-consuming, and the obtained test scores become less reliable.

Practically speaking, only simple computations and verbal associations can be tested satisfactorily by completion items, and the search for such questions too often leads the teacher to concentrate on details. Examples of completion items are shown below.

Examples:

- If a fabric costs $2.00 a yard, one-quarter yard of the fabric would cost _____.
- Grass stains can be removed from cloth by sponging it with _____.

3. True-False Items

The true-false item is the easiest type of item to prepare. True-false items provide a good measure of the degree to which students have learned basic factual material. A wide range of subject matter can be tested by true-false items in a short period of time. They can be quickly and easily scored and scoring reliability is high.

However, true-false items tend to be less discriminating than multiple-choice items and somewhat more subject to ambiguity and misunderstanding. True-false items emphasize recognition over recall and application and may encourage students to develop poor study habits such as rote memorization. True-false items must be based on material that is absolutely true or false and cannot be used to test knowledge of controversial or qualified material. Also it has been determined that for a true-false test to be as reliable as a multiple-choice test, it must contain almost twice as many items. An example of a true-false item is given below.

Example:

T F • A soft-cooked egg should be simmered for ten minutes.

4. Multiple-Choice Items

Compared to other item types, a well-constructed multiple-choice test item has a relatively high ability to discriminate between the students who have and the students who have not obtained the course objectives. A large range of subject matter can be tested by multiple-

choice items in a short period of time. They can be used to measure all levels of knowledge and some intellectual skills. They can be quickly and easily scored and scoring reliability is high.

However, it is often difficult to construct good multiple-choice items, and finding plausible alternatives is sometimes difficult. Multiple-choice items are often difficult for slow readers. An example of a multiple-choice test item is given below.

Example:

- Which document contains detailed information about the title of property?

1 deed	3 bond
2 lease	4 mortgage

5. Matching Questions

The matching question requires relatively little reading time on the part of the students, as compared with other types of items, and a large number of questions can be answered in a short period of time. Scoring is simple and objective.

Unfortunately, it is often difficult to obtain a sufficient amount of related material so that homogeneous lists of premises and responses can be constructed. Furthermore, when the questions and answers can be no more than a word or phrase they usually tend to be somewhat superficial. It is also difficult to eliminate all clues to answers in matching questions. An example of a matching question is shown below.

Example:

- On the line at the left of *each* set of symptoms listed in parts *a* through *e*, write the *number* of the condition, *chosen from the list below*, that would produce that set of symptoms. (5 credits)

Conditions

1. Food poisoning
2. Sunstroke
3. Poison ivy
4. First degree burn
5. Second degree burn
6. Insect bite
7. Shock

_____a. Swelling over joints and discoloration of skin.
_____b. Small, watery blisters which spread.
_____c. Paleness, chills, and a cold perspiration.
_____d. Stomach pains, nausea, and cramps.
_____e. Flushed face, rapid pulse, and dry skin.

C. Measuring Achievement To Assign Grades: Cognitive Learning Outcomes

When constructing test items for the cognitive objectives, the teacher should always remember that a good test item gives the students every possible opportunity to show what they have achieved. A test item will not measure well if it becomes a contest between the test constructor and the test taker, in which one wins and the other loses. A test item is a poor one if it tricks students into giving the wrong answers even though they really know the right answers.

Every test item should be written as clearly, directly, and simply as possible, so that all students understand what is expected of them. The more precisely the test item has been written, the more effectively it will perform the job expected. A difficult test item is not necessarily better than an easy test item, if its difficulty is a result of inappropriateness, high vocabulary level, or a confusing presentation.

It is essential that the test constructor differentiate clearly between the acceptability of the concept being tested and the acceptability of the way in which the test item is presented to students. Following are a number of rules that should be followed by the teacher when preparing essay questions, completion items, true-false items, multiple-choice items, and matching items.

1. Essay Questions

a. Use the essay question to measure objectives that cannot be measured as well with other question types.

One of the common weaknesses in classroom tests is a failure to restrict the essay question to its special purposes, resulting in the less efficient testing of objectives that could be better tested in other ways. The essay question should not be used solely to elicit information. The routine testing of the recall of information is a function that can be performed more effectively by objective-type questions. Consider, for example, the following question:

(Poor)

> Name and locate 5 of the commonest safety hazards that are found in homes.

This question does not call for analysis, judgment, problem solving or any of the abilities toward the measurement of which the essay question can make a special contribution. It affords little opportunity for written expression. Use of the essay form in this instance serves merely to increase rating time, invite rating problems and make the test scores less

meaningful. The teacher may also have considerable difficulty in deciding whether some of the students' borderline answers are really common safety hazards.

In contrast, the following type of question can be defended as having a special contribution which justified the essay approach:

(Improved)

In today's society, women are better accepted in the job market than they were in the 1920's. Do you agree or disagree with this statement? Why?

From their background knowledge and experience, the students must select the pertinent information, evaluate it, decide how best to express it and write the answer. The mental processes the students go through would appear to be basically different from those that the student would use in answering related questions presented in the objective form.

b. Present the students with a clear-cut question that has only one interpretation and correct answer.

Too often an essay question will ask students to write all that they know about a certain topic. A question of this type can have as many different interpretations and correct answers as there are students. Thus, it becomes very difficult to determine the relative achievement of students on the topic covered by the test item. An essay question of this type is shown below.

(Poor)

In a series of paragraphs, explain about renting a house.

In order to be able to compare students' achievement in a certain area, the essay question must be constructed so that it can be interpreted in only one way and has only one specific answer. The question given below is much more specific than the first example and there should be a specific correct answer for each part.

(Improved)

Discuss one advantage and one disadvantage of renting a home for each of the following: (10 credits)

(a) a single person
(b) two single persons (sharing)
(c) a young married couple without children
(d) a young married couple with a small child
(e) a retired couple

A helpful procedure for the teacher to follow is to outline beforehand all the possible answers that students may write in answer to an essay question. Sometimes it becomes immediately apparent that the question is too broad and general and that the scope of the question must be limited. Sometimes the reverse is found to be the case—that there is really not very much the students can give in answer to the question as framed. In such instances, the scope of the question can be broadened appropriately.

c. Use many brief essays rather than one or two extended essays.

There are two important reasons for using brief essays rather than long essays. First, a test composed of many brief essays will permit a wider sampling of the content of the course and will provide a more accurate picture of students' total achievement. Second, the test score will be more reliable. If the essay test contains only a few questions, each calling for a long answer, the students who happen to misinterpret one question will be penalized disproportionately to their real knowledge of the subject. To be reliable, a test must provide for a reasonable number of independent responses. A long essay may be carefully delimited in scope by subdivision into component elements, but these elements are generally too closely related to provide either the breadth of sampling or the higher reliability that can be obtained with an equivalent number of independent brief essays.

d. Indicate clearly in each question the type and length of the answer expected.

In each essay question, the teacher should indicate the type of answer expected by using key words such as "explain," "describe," "outline," "list," "state," etc. The teacher should also indicate how many parts of the answer should have and approximately how long each part of the answer should be. This can be done by providing either the approximate number of words required or the number of credits allocated for each part of the question. Providing this information will help to insure that the students will spend the appropriate amount of time on each question. Consider, for example, the question given below.

Example:

Mr. Moore and five-year-old Deborah have just brought Mrs. Moore home from the hospital with a new baby daughter.

(a) List *four* benefits this family may gain as they care for the new baby. (4 credits)

(b) For each of the family members listed below, list *two* adjustments in living habits that may be necessary because of the new baby. (Use each adjustment only once.) (6 credits)

(1) father
(2) mother
(3) Deborah

In the question given above, the students are told what type of answers are expected by the word "list" in parts a and b. The students are also told that the teacher expects four different answers for part a and six different answers for part b. The number of credits allocated to each part should indicate that each answer can be fairly short. After reading this question, the students should have a clear idea of exactly what must be done and how long should be spent doing it.

2. Completion Items

a. Avoid using ambiguous statements that do not limit the answer to one or two specific words or phrases.

Each completion item should be examined carefully to make sure that it calls for a specific, limited answer. Too often the completion item is so ambiguous that it offers ample opportunity for the students to misinterpret the question or evade it. For example:

(Poor)

Food should always be served to young children in _____.

The teacher should not be surprised to find several different answers such as "small amounts," "dishes," "the morning," "pleasant surroundings," or "preschool." Drawing the line between what does and what does not constitute a correct answer can be a difficult problem. By rewording the item as shown below, the correct answer is limited to one specific word.

(Improved)

When serving food to young children, the size of the portion should always be_____.

b. Do not require more than one completion to be made in an item.

Sometimes completion items are constructed so that two or more blanks are left in the same sentence or paragraph. As the number of blanks increases, the interrelationships and combinations become more complex. Such questions frequently require an excessive amount of time to answer, require careful scoring for part-credit allowance and

introduce the extraneous factor of verbal facility into the measurement picture. In the following item, for example, too many words have been removed and the item does not pose a clear-cut question.

(Poor)

Most (communicable) diseases are caused by (germs) that enter the body through the (nose) and (mouth).

c. The blank should be placed near or at the end of the statement.

If the blank to be filled in by the students comes at the beginning of the statement, the students must begin with a blank, read the statement, and then retrace their steps to decide what should be written in the blank. For example:

(Poor)

_____are the major source of calories in most diets.

The item can be understood more readily by the students if the blank is at the end of the sentence. By the time the students reach the blank space, they should have all the information they need to answer the question. An "improved" version of the preceding item is shown below.

(Improved)

The major source of calories in most diets is_____.

d. Avoid extraneous clues to the correct answer.

Sometimes the grammatical structure of a completion item provides a clue to the correct answer. For example:

(Poor)

In a young child, hyperactivity is usually a symptom of an_____maladjustment.

The use of the word "an" in this item would indicate to alert students that the correct answer must begin with a vowel. This would automatically eliminate such choices as "social," "physical," and "mental" simply because they begin with consonants.

The verb of a completion item can also provide a clue to the correct answer. An example of this type of clue is shown below.

(Poor)

_____are a leading cause of death in young children.

The use of the verb "are" directly after the blank space automatically eliminates such possible answers as pneumonia or tuberculosis. In the

"improved" version shown below, the verb of the item is placed before the blank space and does not provide a clue to the correct answer.

(Improved)

A leading cause of death in young children is _____.

Another type of extraneous clue is the use of short blanks for short words and long blanks for long words. The length of the blank should not indicate the length of the word or phrase to be inserted. Nor should the number of blanks indicate the number of words omitted. In the following item, students are not likely to answer "oxygen" or "nitrogen" since two blanks are provided for the answer.

(Poor)

The gas that makes a cake rise is _____ _____.

While the teacher does not want to discourage alertness in students, the main objective of any achievement test is to measure subject matter knowledge, not test wiseness. The presence of extraneous clues to the correct answer does not further this objective.

3. True-False Items

a. True-false items should always be based on statements that are absolutely true or false, without qualifications or exceptions.

Exceptions to true-false items can cause students to become confused if they are aware of the exceptions. This confusion is due to the fact that the students are not sure of the strictness of the standards of truth that must be applied. The following item is generally true and will be accepted by most students at the superficial level at which it is intended.

(Poor)

T F Bacteria need moisture to live.

Some superior students, however, might think that this is an item designed to test whether they realize that when bacteria form spores, the bacteria do not need moisture to live. On the basis of this exception, the students might mark the item false. Each true-false item should be reviewed carefully in an attempt to determine if there are any possible exceptions that could cause confusion in the minds of the students.

b. Avoid the use of long and involved statements.

Often a test item will contain information that is not essential to answering the item, but that makes it more interesting to the students or more realistic. The addition of nonessential information to an item tends to make the item so long and involved that the students find it difficult to identify the crucial element in the item. For example:

(Poor)

T F Tom and Joan are seriously considering becoming engaged. They should see less of their friends and spend more time alone together in order to become more familiar with each other's habits.

When constructing items, the teacher should always try to make the items as short and easy to read as possible. In doing this, the teacher will ensure that the items are mainly measuring the students' knowledge of required concepts and not their ability to read. An "improved" version of the preceding item is shown below.

(Improved)

T F Before two people become engaged, they should spend time alone together and learn about each other's habits.

c. Avoid negative items whenever possible.

Negative items tend to create confusion by requiring a radical shift in orientation from previous questions. They reduce the test reliability by inducing careless errors due to overlooking the negative. To ask students to conclude that a negative item is false because it is not true that something is not true involves a twist in reasoning that can be extremely confusing.

Usually the instructional objective of the teacher can be achieved just as well and measured much more effectively by the more direct approach. Consider, for example, the following item:

(Poor)

T F Calcium is *not* essential for the growth of bones and teeth.

This same concept can be tested more directly in a positive statement.

(Improved)

T F Calcium is essential for the growth of bones and teeth.

Items containing double negatives are especially likely to be troublesome.

(Poor)

T F A baby does not cry if it does not have a reason.

Again, the same concept can easily be tested in a positive form.

(Improved)

T F A baby cries only when there is a reason.

The only instance in which a negative item might be used is when the concept to be tested is essentially a negative one. In the example

shown below, the concept being tested is normally presented to students in terms of what should not be done rather than what should be done.

(Poor)

T F Iodine should *not* be applied around the eyes.

In such cases it is desirable to emphasize the negative nature of the item by underlining the negative words or by using distinctive type styles.

d. Avoid extraneous clues in the form of specific determiners.

An analysis of the usual true-false examination will reveal that the students who arbitrarily go through the examination and place a "false" next to every item containing an extreme word (such as *always*, *never*, and *all*) and a "true" next to every item containing a qualifying word (such as *usually*, *sometimes*, and *may*) will be right more often than they will be wrong. These words are used frequently by teachers because they are necessary to make the items foolproof—either absolutely true or absolutely false. Students soon begin to base their answer on these keywords rather than on the item as a whole. Examples of items with such extraneous clues are given below.

Examples:

T F Cuticles should *always* be cut.

T F Malnutrition *never* occurs in families that have ample funds to spend for food.

T F Clothes *usually* dry more quickly on a windy day than on a calm day.

This does not mean that qualifying words should never be used in true-false items. It is not possible or desirable to avoid them completely. What should be avoided is the pattern of having certain types of words associated with certain answers. If items containing *always* or *usually* are sometimes true and sometimes false, the students are not provided extraneous clues to the correct answers.

4. Multiple-Choice Items

a. Each item should be based on a single central problem, which is completely stated in the stem of the item.

The item should be worded so that the students will know exactly what is being asked before reading the choices. In the following example, there is nothing in the stem to indicate to the students what the

problem is. Just what does the teacher want to know about pork chops? That can only be determined by reading all of the choices.

(Poor)

Baked pork chops should always be
1 served with applesauce
2 cooked three minutes on each side
3 cooked at 425° F.
4 well done, but not dry

The example shown above is really concerned with four different problems: what should be served with pork chops, how long they should be cooked, at what temperature they should be cooked, and what should be the condition of the meat when it is properly baked. The students are not required to choose the best of the number of choices to a single problem, but rather to decide which of a number of independent true-false statements is more true than others. This is a difficult and confusing task at best.

The best way to avoid this pitfall is to write the item first as a question rather than as an incomplete statement. If the stem of an item cannot be written in the form of a question, then the stem does not contain a completely stated central problem. After the question has been written and the four choices supplied, the question can be rewritten as an incomplete statement if so desired.

The foregoing item can be improved by basing the question on just one problem as shown below. In this item, the students should know exactly what is being asked before reading any of the choices. If the teacher wishes to test the other three topics, separate questions should be written on each topic.

(Improved)

Pork chops should be baked at a temperature between
(1) 250° F and 300° F
(2) 300° F and 350° F
(3) 350° F and 400° F
(4) 400° F and 450° F

b. Each item should be written as clearly and simply as possible.

The ability of an item to discriminate between the students who have obtained an objective and those who have not can be seriously limited when the item is not perfectly clear. The difficulty of an item should be due to the problem involved rather than the manner in which it is presented. In the example shown below, confusion could be caused by

the manner in which the item is worded and the unnecessary use of the word "predominating." By rewording the item and using a simpler vocabulary, the item can be made much clearer as is shown in the "improved" version.

(Poor)

Which is a mixture of many spices, which has tumeric, ginger, and red pepper predominating?

1 curry
2 coriander
3 cumin
4 marjoram

(Improved)

Which seasoning is made mainly of tumeric, ginger and red pepper?

1 curry
2 coriander
3 cumin
4 marjoram

Another common way in which items are made unnecessarily difficult is by presenting too much material in one sentence as shown below. In the "improved" version, the item is divided into two sentences and is much easier to read.

(Poor)

What had Mrs. Brown forgotten to evaluate, when, after purchasing a game, she found that it required six batteries?

1 the operating cost
2 the type of finish
3 the materials of construction
4 the ease of care

(Improved)

After buying a game, Mrs. Brown found that it needed six batteries. What had she forgotten to evaluate?

1 the operating cost
2 the type of finish
3 the materials of construction
4 the ease of care

c. Make all of the choices plausible and attractive to students who lack the information or ability tested by the question.

As far as possible, each choice should be designed specifically to attract students who have certain misconceptions or who tend to make common errors. In the example shown below, choices 1 and 2 are implausible. Most students would probably consider only choices 3 and

4 and have a fifty percent chance of selecting the correct answer, even if they did not know the concept being tested by the question. In the "improved" version, all choices are plausible and the students would have to know something about the concept being tested in order to eliminate any of the choices.

(Poor)

When using elastic in clothing, the elastic should be cut
(1) very slowly
(2) with a razor blade
(3) 3 to 4 inches smaller than the actual measurements
(4) 3 to 4 inches larger than the actual measurements

(Improved)

When using elastic in clothing, the elastic should be cut
(1) 5 to 6 inches smaller than the actual measurements
(2) 3 to 4 inches smaller than the actual measurements
(3) the same length as the actual measurements
(4) 3 to 4 inches larger than the actual measurements

d. Each item should have one and only one correct answer.

While this requirement seems obvious, it is not always fulfilled. In the example shown below, choice 4 was keyed as the correct answer. But after examining the other choices, it can be seen that choices 2 and 3 are also correct.

(Poor)

Before storing a cutting board, the cook should
1 wash it in a dishwasher
2 brush off the crumbs
3 wipe it off with a damp sponge
4 sanitize it with a mild chlorine solution

Most often this problem occurs when the item writer fails to consider the full implications of each choice.

e. Avoid negative questions if possible.

Negative statements in multiple-choice items lead to the same difficulties and confusions as in true-false items. In the case of multiple-choice items, special care must be taken to avoid the double negative. To the item writer, a combination of negatives in the stem and in the choices may not be especially troublesome. The students, however, are likely to be confused by an item containing a double negative. As can be seen from the two examples below, a positively stated item is much clearer than an item containing a double negative.

(Poor)

When a child feels warm and complains of a headache, what should a babysitter *not* do?

1 The babysitter should not call the child's parents.
2 The babysitter should not have the child lie down.
3 The babysitter should not give the child aspirin.
4 The babysitter should not assure the child that he will feel better.

(Improved)

When a child feels warm and complains of a headache, the babysitter should

1 give the child an aspirin
2 call her mother for help
3 call the child's parents for instructions
4 contact her family doctor

f. Avoid extraneous clues.

The teacher must be very careful to avoid extraneous clues that would help students to choose the correct answer or eliminate an incorrect choice. Students' scores on a test will be a less accurate measure of what they have achieved if an extraneous clue helps them to choose the correct answer or eliminate an incorrect choice. Such clues may assume a variety of forms:

- If keywords are found in both the stem of an item and one of the choices, that choice is usually the correct answer.
- If one choice is considerably longer than the other choices, that choice is usually the correct answer.
- If a choice is not grammatically consistent with the stem of an item, that choice is usually not the correct answer.
- The choice "none of these" should be used with caution. If the choice "none of these" is used, it should be used in several items that are grouped together and it should occasionally be the correct answer. In many classroom tests, however, "none of these" is used only because the teacher cannot think of a good fourth choice and it is rarely the correct answer. Once students realize that "none of these" is rarely the correct answer, they will automatically eliminate it for a reason which is not related to the concept being tested.

5. Matching Items

a. Each matching question should contain homogeneous premises and homogeneous responses.

The basic problem in preparing matching questions is to prepare a question that has homogeneous premises and homogeneous responses. If a matching question does not contain homogeneous premises and homogeneous responses, it will be impossible to: (1) prepare directions that clearly explain the basis upon which the items are to be matched, and (2) eliminate all extraneous clues that would enable students without the proper subject matter knowledge to make correct matches. An example of a matching question that does not have homogeneous premises and responses is shown below.

(Poor)

Match the phrase in Column B to the word in Column A with which it is most closely associated. (5 credits)

Column A	Column B
_____a. Crepe Suzette	1) used to cut up vegetables
_____b. No. 8 scoop	2) milk with fat removed
_____c. Skimmed	3) used to serve fried chicken
_____d. French knife	4) French pancake served aflame
_____e. Brown Betty	5) milk before any processing
	6) a baked pudding of fruit
	7) used to serve mashed potatoes

The best way to determine if a matching question contains homogeneous premises and homogeneous responses is to ask yourself the following question about each column: "Can I, with one or two words, or a short phrase, specifically describe what is contained in each column?" If the answer is *no*, then the matching question does not contain homogeneous premises and homogeneous responses. For example, in the question given above, column A contains the names of two types of desserts (a, e), two types of kitchen utensils (b, d), and one type of milk (c). Column B contains three phrases which describe the uses of kitchen utensils (1, 3, 7), two phrases which describe types of milk (2, 5), and two phrases which describe types of desserts (4, 6).

Since the question given above does not contain homogeneous premises and homogeneous responses, it is impossible to prepare a set of directions that gives the basis upon which the items are to be matched and also to eliminate extraneous clues. The matching question given above actually contains the basis for three questions, one on types of desserts, one on types of kitchen utensils, and one on types of milk. The improved question shown below was developed from the section on types of desserts.

(Improved)

For each description given in parts a through e, select the name of the dessert, chosen from the list below, that is being described. (5 credits)

Names of Desserts

(1) Baked Alaska
(2) Bavarian
(3) Brown Betty
(4) Cherries Jubilee
(5) Compote
(6) Crepe Suzette
(7) Peach Melba

_____a. A baked pudding of fruit
_____b. Fruits stewed in a syrup
_____c. A French pancake served aflame
_____d. Whipped gelatin and cream
_____e. A cake covered with ice cream and meringue

b. Keep the lists of premises and responses relatively short.

The matching lists should be kept short so that they can be easily handled by students. Some matching questions in school examinations have as many as 20 or 30 items in each list. If the students are to compare each of 30 premises with each of 30 responses, they are required to make a total of 30 × 30 or 900 comparisons. Obviously, this is too long and tedious a task to impose upon them. Economy of testing time, generally a strong advantage of the matching question, is no longer achieved.

There are other disadvantages as well. The strong possibility that students may overlook the correct answer, buried as it is in a mass of other responses, tends to reduce validity. From the teacher's point of view, it is difficult to maintain homogeneity in a long list of items. Further, a long list of matching items usually means excessive concentration on one aspect of achievement, preventing proper distribution of emphasis on other aspects.

A popular size in teacher examinations is the 10-item matching unit, which seems to work fairly well. Probably the optimum size from the testing viewpoint, however, would be the matching question containing 5 to 8 items.

c. Provide for extra responses to reduce guessing.

One of the advantages of the matching question is that it tends to minimize the guessing factor. In a 10-item matching question, for

example, if students do not know the answer to one of the items, their chances of guessing the right response are rather small. In many teacher examinations, however, the practice is to offer a list of premises and a list of responses of equal size. As a result, if a student knows 8 items of the 10, the student's chance of guessing correctly the responses to each of the remaining 2 items becomes 1 out of 2. If the student knows 9 of the 10, then the student can answer the remaining item correctly merely by a process of elimination. It is therefore generally considered good practice to have the number of responses exceed the number of premises. If there are 5 items, 7 responses might be offered. If there are 7 items, 10 responses might be offered.

D. Measuring Achievement To Assign Grades: Psychomotor Learning Outcomes

1. Introduction

As mentioned previously, student progress toward the achievement of the psychomotor objectives should be measured by performance tests. A performance test involves the careful and systematic observation and evaluation of either the procedures executed by students when performing a job (process evaluation), the product produced by these procedures (product evaluation), or a combination of both.

Evaluating the characteristics of the product is perhaps the more efficient of the two kinds of evaluation for two reasons. First, individual student observation is not necessary; individual products can be evaluated after the class period, thus requiring less class time. Second, less teacher time is involved, since evaluating a product may take only a few minutes, whereas observing and evaluating the process involved may be quite lengthy, depending on how much time is involved in making the product.

When evaluating the procedures involved in student performance (process evaluation), the teacher must observe each student individually. Although this can be a time-consuming process, on some occasions it may be impossible for the teacher to infer the students' ability to perform certain procedures by evaluating only the product. For example, it would be impossible to determine how safely or how efficiently a student had performed a job by evaluating only the product of the job. In these instances, the teacher must actually observe each student's performance in order to determine if the student has satisfactorily performed these procedures.

In summary, the only procedures that should be actually observed and evaluated by the teacher are those that cannot be inferred from evaluating the product.

When evaluating procedures and the characteristics of a product, the teacher may use either a checklist, a rating scale, or both. When a teacher is concerned mainly with whether certain procedures and characteristics of a product are either present or absent and is not concerned with the degree to which they are present, then a checklist should be used. When the teacher is concerned with the degree to which certain behaviors and characteristics of a product are present, then a rating scale should be used. The teacher may decide to use both techniques when constructing performance tests.

Checklists consist of a vertical listing of the characteristics of a product or a series of procedures. The procedures should be listed in the order of occurrence. For each item listed, the observer indicates with either a checkmark or a yes-no response whether or not the required behavior had occurred or the required characteristic was present. For a checklist, the raw score would be equal to the total number of checkmarks or the total number of "yeses."

Rating scales also consist of a vertical listing of the characteristics of a product or a series of procedures. In a rating scale, however, a continuum of merit or quality is established for each item listed. The continuum is usually divided into three to five positions which are listed horizontally. Each position is identified by a number, and represents a certain degree of merit or quality. The observer then rates each item by choosing the position that corresponds most closely to the degree of merit or quality observed. In the scoring of the rating scale, the raw score equals the sum of the ratings for all of the items on the rating scale.

2. Performance Test Construction Procedures

When developing performance tests for a specific objective, the teacher should first list all of the procedures that the students would be expected to perform, as well as the supplies and equipment that the students would be expected to use in accomplishing that objective.

The teacher can then prepare a test exercise that would require the students to perform the procedures and use the supplies and equipment that were previously listed. If desired, the production of a particular product may be designated as a part of the test exercise.

The next step is to decide which procedures and/or characteristics of the product are to be evaluated. As mentioned above, the only procedures that should be evaluated are those that cannot be inferred from evaluating the product. The teacher should now prepare the checklist and/or rating scale that will be used to evaluate the selected procedures and/or the characteristics of the product.

Following below is a sample psychomotor objective which can be evaluated by means of a performance test. This objective will be used to illustrate the steps outlined above for constructing performance tests.

It should be noted here that in constructing the performance tests for the sample objective below, some of the procedures (6b, 6d, 6e) included in the first list of procedures are not included in the checklist of procedures or the rating of procedures. These procedures can be evaluated by observing the characteristics of the product and need not be observed directly.

It should also be noted that, in the sample below, a checklist, a three-point rating scale, and a five-point rating scale are used to evaluate student performance. Because of the three different methods of recording the observer ratings, the raw scores for each method must be converted before they can be combined to obtain a total score for the performance test exercise. The method used to convert the raw scores is shown in the sample below. After the converted scores have been obtained, they can be added together to obtain a total score for the entire performance test exercise. For the following test exercise, the criterion student performance is a total score of at least 75% on the performance test exercise, as stated in the sample objective.

3. Sample Performance Test Exercise

a. Behavioral Objective

Given a recipe for preparing scrambled eggs and following the procedures discussed in class, the student will be able to obtain a score of at least 75% on a performance test exercise that requires the student to select the necessary supplies and equipment, prepare the recipe of scrambled eggs, and clean up after the scrambled eggs have been prepared.

b. List of Procedures

1. The student makes the appropriate preparations for work in the foods laboratory (ties back hair, washes hands, puts apron on, etc.).
2. The student reads the recipe (generally).
3. The student selects and gathers the necessary supplies and equipment from those available in the foods laboratory.
4. The student prepares the working areas:
   a. sink
   b. preparation center
   c. range
5. The student rereads the recipe (specifically).

6. The student follows the recipe:
   a. uses the appropriate measuring procedures
   b. measures the ingredients accurately
   c. handles the ingredients properly
   d. combines the ingredients properly
   e. cooks the product according to the recipe
7. The student washes the dishes.
8. The student cleans the working areas:
   a. sink
   b. preparation center
   c. range
9. The student returns the supplies and equipment to the proper storage areas.
10. The student follows the appropriate kitchen management procedures (gathers supplies efficiently, prepares recipe efficiently, etc.).
11. The student follows the appropriate sanitary procedures (does not taste food with fingers, cleans up spills immediately, etc.).
12. The student observes the appropriate safety precautions (handle of pan turned in on stove, stove turned off immediately after cooking is finished, etc.).

c. List of Supplies and Equipment

measuring cups	salt
measuring spoons	pepper
mixing bowl	butter
fork	apron
fry pan	hot pads
recipe for scrambled eggs	towels
beating utensils	soap
eggs	wash cloth
milk	

—and other necessary supplies and equipment available in a foods laboratory

d. Performance Test Exercise

Using the recipe for scrambled eggs provided by your teacher and following the procedures discussed in class, select the appropriate supplies and equipment from those available in the foods laboratory, prepare one recipe of scrambled eggs, and clean up after the scrambled eggs have been prepared.

e. Evaluation of Performance (Process and Product)

- Process Evaluation

For the process evaluation of the above behavioral objective, two techniques will be used: a checklist and a rating scale.

Checklist of Procedures

		Performed no = 0 yes = 1
1. The student made the appropriate personal preparations for work in the foods laboratory (tied back hair, washed hands, put apron on, etc.).	1.	
2. The student read the recipe (generally).	2.	
3. The student selected and gathered the necessary supplies and equipment.	3.	
4. The student prepared the working areas:		
a. sink	4a.	
b. preparation center	4b.	
c. range	4c.	
5. The student reread the recipe (specifically).	5.	
6. The student followed the recipe:		
a. used the appropriate measuring procedures	6a.	
b. handled the ingredients properly	6b.	
7. The student washed the dishes.	7.	
8. The student cleaned the working areas:		
a. sink	8a.	
b. preparation center	8b.	
c. range	8c.	
9. The student returned the supplies and equipment to the proper storage areas.	9.	

raw score

maximum raw score = 14

converted score = $\frac{\text{raw score}}{1}$*

Rating of Procedures

	never -1-	seldom -2-	some-times -3-	usually -4-	always -5-
1. Kitchen Management					
2. Sanitary Procedures					
3. Safety Precautions					

raw score

maximum raw score = 15

converted score = $\frac{\text{raw score *}}{5}$

- Product Evaluation

The product should be divided onto three plates, so that it can be rated by three judges: the cook, the teacher, and a visitor. The following rating scale may be used by each of the three raters; thus three copies are needed for each student.

Rating of Product: Scrambled Eggs

	Poor -1- Fair -2- Good -3-	
Appearance	1. dull .....shiny	1.
	2. fine pieces .....large pieces	2.
Moisture Content	3. dry or watery.....slightly moist	3.
Texture	4. coarse .....smooth	4.
Lightness	white and yolk white and yolk 5. not well blended .....well blended	5.
	6. compact and heavy .....fairly light	6.
Tenderness	7. rubbery or tough .....tender	7.
Taste	8. stale, flat, or salty .....well seasoned	8.
Flavor	9. raw or burned .....well cooked	9.

$$\text{converted score} = \frac{\text{raw score}}{\text{maximum raw score} = 27}$$

$$\text{converted score} = \frac{\text{raw score*}}{3}$$

*In each case, the converted score equals the raw score divided by the maximum number of points the student can obtain on an item for that evaluation technique.

III. PREPARING INSTRUMENTS USED TO EVALUATE AND PLAN INSTRUCTION

A. Evaluating and Planning Instruction: Cognitive Learning Outcomes

Section II-C of this booklet contains suggestions for preparing test items that can be used to measure student achievement of the cognitive outcomes for the purpose of assigning grades. These same test items also can be used by the teacher to evaluate and plan instruction. For example, if several students receive poor test scores on items testing the knowledge of certain facts or the understanding of certain concepts, the teacher may decide that these facts or concepts need reteaching and plan further instruction accordingly.

At other times, the teacher may wish to measure students' achievement of the cognitive outcomes solely for the purpose of planning instruction. For example, the teacher may administer a group of test items prior to instruction in order to determine how far students have progressed toward achieving certain cognitive outcomes. The test results would then be used only to plan the most appropriate instructional program for these students.

Thus, the same test items and test results that are used to assign student grades also can be used to evaluate and plan instruction; only the intended purpose would change.

B. Evaluating and Planning Instruction: Psychomotor Learning Outcomes

Section II-D of this booklet contains suggestions for preparing performance tests that can be used to measure student achievement of the psychomotor outcomes for the purpose of assigning grades. The same performance tests can be used by the teacher to evaluate and plan instruction. For example, if several students obtain poor scores on the performance of certain skills, the teacher may decide that these students need additional practice in order to obtain mastery of these skills and plan instruction accordingly.

The teacher may decide to have the students rate each other's performance. These peer ratings could not be used to assign student grades because of potential bias and unreliability, but the results may be reliable enough to be used to evaluate and plan instruction. Indirectly, another objective may be achieved: in rating their peers, the students may become more keenly aware of the procedures and criteria, and thus enhance their own acquisition of the observed skills.

The teacher may also decide to have the students rate their own performance. Again, the results would not be reliable enough to assign grades, but as a result of this self-evaluation, the student and teacher may plan future instructional experiences appropriate to the individual student's needs.

Thus, the same performance tests that are used to assign student grades can also be used to evaluate and plan instruction; only the intended use would be changed. Whenever students are rating themselves or each other, the teacher should be sure to inform them that the test results are *not* being used to assign grades. This will help to insure that their ratings will be a reliable indication of student performance.

Two examples of performance tests that could be used to plan and evaluate instruction are shown below. These performance tests were prepared according to the procedures discussed in Section II-D. It should be noted that since the second example was designed as a self-inventory, the statements to be rated were phrased in the first person.

Storyreading Effectiveness

Directions: Check the column that best describes the way the reader reads the story aloud to the children. Score the performance 1–3 according to the procedures followed.

	never -1-	some-times -2-	always -3-
1. Reads with expression			
2. Speaks distinctly			
3. Maintains children's attention			
4. Holds book so that all children can see			
5. Looks at children while reading			
6. Provides time for questions and discussion			

total score

(total possible points = 18 points)

Rating:

15–18 points—Excellent storyreading
10–14 points—Fair storyreading
6–9 points—Poor storyreading

(*Note:* This rating scale may be scored by the teacher, another observer, or another student.)

Classroom Manners: A Self-Inventory

Directions: Rate yourself from 1 to 5 to describe what you do to help maintain a pleasant classroom atmosphere. Total your responses below. Be honest! Results may be used to formulate new procedures in areas where improvement is needed.

	never -1-	seldom -2-	sometimes -3-	usually -4-	always -5-
1. I get to class on time					
2. I enter quickly and quietly					
3. I am orderly during class					
4. I get ready for work quickly					
5. I am thoughtful and considerate					
6. I keep my materials where they belong					

total score

(total possible score = 30 points)

Self-Rating:

22–30 points—very good manners
13–21 points—manners need some improvement
6–12 points—manners need very much improvement

C. Evaluating and Planning Instruction: Affective Learning Outcomes

1. Introduction

As mentioned previously, the courses of study in home economics contain affective outcomes which involve " . . . changes in interest, attitudes and values, and the development of appreciations and adequate adjustments." When planning and evaluating instruction, the teacher should measure student growth relative to the affective outcomes as well as of the cognitive and psychomotor outcomes.

Some of the instruments that may be used to measure student achievement of the affective outcomes are questionnaires, inventories, logs, diaries, and anecdotal records. The validity and reliability of the information obtained by these instruments is closely related to the quality of the instruments themselves. Poorly constructed instruments are likely to produce inaccurate results. Therefore, it is important that these instruments be as well constructed as possible. The following sections will contain a brief discussion of the uses and limitations of each type of instrument, guidelines that should be followed when constructing each type of instrument, and illustrative examples.

2. Questionnaires and Inventories

a. Use of Questionnaires and Inventories

The use of questionnaires and inventories should be limited to situations where the desired information cannot be obtained more easily in another manner, such as by an interview. Questionnaires and inventories are used to gather information about the interests, needs, concerns, opinions, and judgments of students as well as to obtain factual information about students' backgrounds, family, and home conditions. They may be used to obtain information about students from not only the students themselves, but also from the students' peers and parents and other adults.

b. Limitations of Questionnaires and Inventories

Whenever interpreting the information obtained by questionnaires and inventories, the teacher should always remember that this information may be subject to some unreliability (inconsistent responses upon retesting at later times). This unreliability is partially due to the fact that students' interests, attitudes, and values are constantly changing, particularly at the junior and senior high school level.

Other factors that may cause the information obtained by these instruments to be unreliable are the respondents' degree of awareness of

their own attitudes, interests, opinions, or knowledge, their ability to verbalize them, and the respondents' willingness to tell the truth. The respondents' willingness to tell the truth can be increased by instructing them, whenever possible, to respond anonymously. This practice will help to encourage the students to answer more frankly and will enable the teacher to plan instruction more accurately.

c. Guidelines for Preparing Questionnaires and Inventories

The formats of questionnaires and inventories are somewhat different. Questionnaires usually ask respondents to provide information in several related areas. The respondents may be asked to provide the information by checking the correct answer to a question, or by writing one or two words, a short sentence, or even a paragraph. An inventory usually covers only one area of consideration and the respondents provide the requested information by checking their answer to the question being asked. The guidelines given below should be followed when constructing questionnaires and inventories:

- Clearly define the purposes of questionnaires and inventories so that all necessary items will be included and all unnecessary items excluded.
- Limit the total length of questionnaires and inventories and the length of each item so as to insure that the students will carefully answer each item.
- Use understandable vocabulary.
- Limit each item to one area of consideration to help avoid difficulty in interpreting responses.
- Provide sufficient response categories so as to include all possible responses.
- Include "cross-checking" questions to determine consistency of responses.
- Group together questions that deal with the same area. Within each area, group together questions that use the same type of response format (true-false, multiple-choice, fill-in-the-blank, etc.).
- Design the format attractively (well spaced, well printed) so as to motivate the respondent to completely answer all questions.
- Design the groupings and response categories so as to provide easy summary and interpretation of results.

d. Examples of a Questionnaire and an Inventory

Following are examples of a questionnaire and an inventory.

Questionnaire — Clothing and Textiles

Name _____ Male_____ Female_____

School _____ Grade_____

1. List previous school courses in clothing and textiles with a brief description of course content.

_____

_____

_____

_____

_____

2. List any out-of-school courses or clothing construction experiences you have had (such as 4-H) with a brief description of what you did.

_____

_____

_____

_____

_____

3. List the garments or clothing and textile products you have made.

_____

_____

_____

4. To what extent do you feel skilled in the following areas:

	No Skill	Fairly Skilled	Very Skilled
a. selecting a pattern			
b. selecting fabrics			
c. purchasing fabric for a pattern			
d. measuring and fitting for a pattern			
e. laying and cutting a pattern			
f. using a sewing machine			
g. sewing in a zipper			
h. making buttonholes			
i. ironing and pressing			
j. hemming garments			
k. altering garments			
l. repairing garments			

Home Decorating Interest Inventory

Directions: The purpose of this inventory is to determine your needs and interests in the area of home decorating so that the appropriate learning experiences may be planned. Check the column that best describes your level of interest.

Are you interested in:	very interested	somewhat interested	not at all interested
1. The use of principles of design in coordinating a room?			
2. Different styles of furniture?			
3. Considerations in wise selection of furniture?			
4. New ideas of furniture arrangement?			
5. Making the home decoration dollar stretch?			
6. Different qualities of furniture?			
7. New ideas of finishing and refinishing furniture?			
8. Selecting accessories for the home?			
9. Creating accessories for the home?			
10. Taking a field trip to the following places:			
a. furniture store			
b. furniture manufacturer			
c. flea market			
d. auction			
e. antique shop			
f. crafts shop			

3. Logs and Diaries

a. Uses of Logs and Diaries

Logs and diaries can both be used to record information about student activities, needs, feelings, and attitudes that cannot be obtained by the more objective test procedures. The information recorded in logs and diaries can be used as a basis for evaluating student achievement of the affective outcomes and for planning future instruction. The same information can also serve as a basis for discussion during teacher-student conferences and classroom instruction.

There is no sharp distinction between logs and diaries. The same type of behavior can be recorded in both. The main difference is that in a log, recordings are made only when the behavior being observed occurs, while in a diary, recordings are made on a daily basis over a certain period of time.

Logs and diaries are especially effective in the areas of family relations, child development, and management. Logs can be used to record observations that students might make while babysitting or observing children in nursery school or kindergarten. Logs and diaries can both be used to record information that could be used to review management practices or evaluate home responsibilities. They both can be used to record information about interpersonal relationships that could provide students with insights into their own behavior.

b. Limitations of Logs and Diaries

The information recorded in logs and diaries may be influenced by several factors. A major limitation is that students may not wish to keep a revealing account of personal feelings and data. In fact, those students who need help the most are the ones least likely to record personal feelings or data. Also, a recording can easily be biased by the mood that students are in while making the recording. A happy student and a depressed student could make two entirely different recordings of the same observation. The information recorded in logs and diaries may also be influenced by the degree of insight that students have into their behavior and their ability to document it.

c. Guidelines for Constructing Logs and Diaries

The guidelines given below should be followed when constructing logs and diaries.

- Determine the purpose of the log or diary, i.e., how will the information recorded in the log or diary be used by the pupil or teacher? For example, a teacher could use the information recorded in a log or diary to help students to:

— become aware of how they spend their time
— become aware of their values, attitudes, and appreciations
— learn why preschool children behave as they do
— evaluate their home responsibilities
— evaluate their interpersonal relationships with other people

- After the purpose of the log or diary has been determined, the teacher must next decide which specific information must be recorded by the students in order for the log or diary to serve its intended purpose.
- The teacher should now prepare specific directions that will tell the students exactly what type of information should be recorded. This will insure that all relevant information will be included and all irrelevant information excluded.
- Whenever students are recording information about the occurrence of an event, the students should be instructed to record information either during or immediately after the occurrence of the event. The longer the students wait to make their recording, the less accurate the recording will tend to be.

d. Examples of a Log and a Diary

Log: Nursery School Observation

Purpose: To observe and record the behavior of two-year-old children in a nursery school setting in order to learn to recognize different patterns of social interaction.

Directions: On the first day of your observations at the nursery school, select one child who does and one who does not appear to be associating freely with his or her peers and adults. You will observe these two children for one-half hour twice a week during the next 6 weeks.

On the form provided for each child, record all significant behaviors exhibited by these two children relative to social development, and add any comments that you might wish to make. Be sure to include the setting in which the recorded incidents took place.

Observation Log

Name of observer _____

Nursery school _____

Name of child observed _____

Date, time, and setting of observation	Events Observed	Comments

Diary: Homemaking Tasks

Purpose: To record the homemaking tasks performed at home for 1 week to determine the type and frequency of these tasks.

Directions: Keep a diary for 1 week of the homemaking tasks you perform each day at home. At the end of the week, total the frequencies for each type of task and indicate the total number of tasks performed.

Monday

_____

_____

Tuesday

_____

_____

Wednesday

_____

_____

Thursday

_____

_____

Friday

_____

_____

Saturday

_____

_____

Sunday

_____

_____

Type of Task	Frequency
1. Cleaning	
2. Cooking	
3. Child care	
4. Laundry	
5. Ironing	
6. Sewing	
7. Mending	
8. Shopping	
9. Yard work	
10. Miscellaneous	

Total Number of Tasks_____

(*Note:* On an actual form, more space should be provided for daily records.)

4. Anecdotal Records

a. Use of Anecdotal Records

Anecdotal records are factual reports of exactly what a person said or did in certain concrete situations. Anecdotal records should be limited to the areas of social, personal, and emotional adjustment and growth. Data on physical status, academic achievement, and study skills can be obtained more efficiently by other means. Observations in these areas should be included in anecdotal records only when the observations might have a marked effect on a student's personality.

An anecdotal record is usually assembled over a long period of time. It helps the teacher to gain an understanding of a student's personality and needs. Once the teacher has gained this understanding, the teacher is better able to plan instruction and develop a classroom atmosphere that will help the student to develop a healthy personality.

Anecdotal records can serve as a basis for individual conferences with students and parents. They may also serve as a cross-check on a student's recordings in a log or diary.

b. Limitations of Anecdotal Records

One limitation of anecdotal records is the time required to prepare them. For this reason, a teacher using anecdotal records usually studies only two or three students at one time.

A second limitation of anecdotal records is that they may be based upon inaccurate observations. Many people have the tendency to see or hear what they expect, rather than what may actually occur. In addition, the incident being observed may not always be recorded objectively. Many people find it difficult to record a purely factual account of an incident. They tend to inject their opinions and feelings into the recording without labeling them as such.

Another limitation of anecdotal records is that they may not contain an adequate sampling of student behavior. A student's record may contain summaries of mainly favorable or unfavorable incidents which could lead to an invalid conclusion about a student's behavior or personality.

c. Guidelines for Preparing and Using Anecdotal Records

The guidelines given below should be followed when preparing and using anecdotal records.

- Study a limited number of students at one time. This will help to keep the teacher from being swamped by a mass of detailed observations.
- Record only those incidents that will contribute significantly to the solution of the problem being studied. If there is any question

whether an incident is significant or not, the teacher should record the incident and then discard it later if it is found to be unrelated to the problem being studied.

- Record both favorable and unfavorable incidents that are significant. A true picture of a student's behavior will not emerge if only favorable or unfavorable incidents are unrecorded.
- Record incidents both in and out of class. Each student should be observed in as many different situations and activities as possible in order to obtain a representative sampling of a student's behavior.
- Record each incident as soon after it occurs as possible. The longer a teacher waits to make a report, the more distorted the report will tend to be.
- Each report should contain some statement of the situation in which the incident occurred so that the report can be properly interpreted at a later date.
- Each report should be as brief and objective as possible. The teacher should not include personal reactions and interpretations in the factual report of the incident. If the teacher wishes to include an interpretation or recommendations for the specific incident, they should be recorded separately and labeled as such.
- Generalizations and judgments about student behavior should not be made until a fairly large number of reports are available. An erroneous picture of pupil behavior may occur if generalizations and judgments are based on limited information.
- Record the report of each incident on a separate index card. The following information should be recorded for each incident:
  - — Name of student
  - — Name of observer
  - — Date of observation
  - — Setting or background of incident
  - — Factual report of incident
  - — Interpretation of incident and recommendations (optional)

d. Example of an Anecdotal Report Form

Anecdotal Record Form

Name of Student _____

School _____ Grade____

Date of Observation _____

Observer's Signature_____

Setting:

Incident:

(front of index card)

Interpretation:

Recommendations:

(reverse side of index card)

(*Note:* Additional cards may be used for multiple anecdotal records.)

References

Ahmann, J. Stanley and Marvin D. Glock. *Evaluating Pupil Growth*. Boston: Allyn and Bacon, Inc., 1971.

Bloom, Benjamin S., ed. *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain*. New York: David McKay Company, Inc., 1956.

Cross, Aleene. *Home Economics Evaluation*. Columbus, Ohio: Charles E. Merrill Publishing Co., 1973.

Greene, Harry A., Albert N. Jorgensen, and J. Raymond Gerberich. *Measurement and Evaluation in the Elementary School*. New York: Longmans, Green and Co., 1953.

Schwartz, Alfred, Stuart C. Tudeman, and Donald G. Wallace. *Evaluating Student Progress in the Secondary School*. New York: Longmans, Green and Co., 1957.

Thorndike, Robert L., ed. *Educational Measurement*. Washington, D.C.: American Council on Education, 1971.

Tinkelman, Sherman N. *Improving the Classroom Test, A Manual of Test Construction for the Classroom Teacher*. Albany, New York: State Education Department, 1968.

Wrightstone, J. Wayne, Joseph Justman, and Irving Robbins. *Evaluation in Modern Education*. New York: American Book Company, 1956.